

The short and very short forms of the Children's Behavior Questionnaire (CBQ) in a community sample of preschoolers

Post-print version

Osa, N. de la, Granero, R., Penelo, E., Doménech, J.M., & Ezpeleta, L. (2014). The short and very short forms of the Children's Behavior Questionnaire (CBQ) in community sample of preschoolers. *Assessment*, *21*, 463-476. doi: 10.1177/1073191113508809

Abstract

The aim was to test the internal structure of scores on the short and very short forms of the Children's Behavior Questionnaire (CBQ) scale and to study the relationship between the dimensions derived and external variables previously related to extreme temperament in a Spanish community sample. The sample was comprised of 622 three-year-old children participating in a longitudinal study. Data were obtained from parents and teachers through a semi-structured diagnostic interview and questionnaires evaluating children's characteristics and psychological states. Results showed a three-factor structure and moderate reliability of the scale scores for both the short and very short forms. Associations were found between the Surgency/Extraversion dimension and ADHD and externalizing problems, between Negative Affect and internalizing and emotional problems, and between Effortful Control and attention, externalizing, and social problems and other executive function measures. Salient temperamental characteristics predicted psychopathological disorders and impairment at ages 3 and 4. The short forms of the CBQ provide reliable and valid scores for assessing temperamental characteristics in the preschool years.

Keywords: Assessment, Children's Behavior Questionnaire, Factor analysis, Preschoolers, Temperament

Introduction

Temperament has been defined by Rothbart, Ahadi, Hershey, and Fisher (2001) as individual differences in reactivity and self-regulation that are constitutionally-based. There is widespread consensus that temperament traits are the core of future personality characteristics (Tackett, Balsis, Oltmanns, & Krueger, 2009). Among the proposed models, to further understanding of the relationship between temperament traits and future personality, the vulnerability/predisposition model states that early extreme temperament may place a child at a greater risk of developing a particular form of psychopathology (Tackett, 2006). Extreme temperament has been identified as a risk factor for subsequent cognitive functioning (Blair, 2002; Healey, Brodzinsky, Bernstein, Rabinovitz, & Halperin, 2010; Wolfe & Bell, 2007), psychopathology (Cipriano & Stifter, 2010; Muris & Ollendick, 2005), and overall functioning (Healey et al., 2010). Longitudinal studies have shown that children's temperament traits are relatively stable across adolescence and adulthood (Asendorpf, Denissen, & van Aken, 2008; Caspi et al., 2003; Putnam, Rothbart, & Garstein, 2008). Also, longitudinal research under a spectrum model conceptualization suggests that some disorders could lay on a similar dimension or continuum with personality characteristics. Some authors suggest that early-measured extreme temperamental characteristics could reveal the early presence of the disorders themselves (Egger & Angold, 2006). Such evidence of a relationship between children's temperament and later developmental problems have led to the search for more efficient instruments with a developmental perspective that can be used at an early age in order to better understand the relationship between temperament and psychopathology and ideally to predict and develop effective preventive measures for personality problems at later ages.

The Children's Behavior Questionnaire (CBQ) (Rothbart et al., 2001) is the tool used the most frequently to assess temperament in children aged 3 to 7 (Putnam & Stifter, 2008).

CBQ, the Infant Behavior Questionnaire-Revised (IBQ-R; Garstein & Rothbart, 2003), and the Early Childhood Behavior Questionnaire (ECBQ; Putnam, Garstein, & Rothbart, 2006), make up a family of questionnaires for fine-grained longitudinal research on temperament. The CBQ is based on the comprehensive view of temperament as constitutionally-based individual differences in reactivity, affect, and self-regulation influenced by heredity and experience (Rothbart & Derryberry, 1981). The standard form of the CBQ (Rothbart et al., 2001) consists of 195 items on 15 scales with 12 to 14 items each targeted at parents of children aged 3 to 7 years old. It is a rationally-generated instrument for assessing theoretically derived core constructs of temperament dimensions. These constructs are broken down into subconstructs, and items are then generated to reflect these subconstructs. Factor analyses of the CBQ have consistently identified three broad temperament traits (Ahadi, Rothbart, & Ye, 1993; Goldsmith, Buss, & Lemery, 1997; Kochanska, Devet, Goldman, Murray, & Putnam, 1994; Rothbart et al., 2001), namely Surgency/Extraversion (SE), involving the tendency to act with impulsive and active behavior which includes positive affect; Negative Affect (NA), referring to the predisposition to experience negative feelings and difficulty being soothed; and Effortful Control (EC), which involves self-regulation, including voluntary regulation of attention and behavior (Rothbart et al., 2001). The invariance in the structure of child temperament has been investigated using the standard form of the CBQ in China (Ahadi et al., 1993), Japan (Rothbart et al., 2001) and the Netherlands (Sleddens, Kremers, Candel, De Vries, & Thijs, 2011). Former studies showed a quite stable invariance in structure of the temperament scale scores. Many other studies have used the instrument with specific populations, such as children with Down's syndrome, stuttering, and voice disorders, and those who were born prematurely (Eggers, De Nil, & Van den Bergh, 2009; Nygaard, Smith, & Torgersen, 2002), as well as to study temperament

clusters in preschoolers (Aksan et al., 1999) and genetic and environmental influences on temperament (Goldsmith et al., 1997).

The publication of short forms of the standard CBQ (Putnam & Rothbart, 2006) was designed to benefit researchers who, like us, need a fine-grained temperament measure but consider the standard form inappropriate due to time limitations. The 94-item short form (CBQ-SF) retains six to eight items of the standard form in each subscale. The decision on which items to include was based on the item content and within-scale factor analysis of the standard scales, as well as the appropriateness and usefulness of the items across the age range of the questionnaire. The resulting CBQ-SF questionnaire assesses the same dimensions as the standard form.

To our knowledge, the CBQ-SF scale scores have only been validated in the US using data extracted from the standard form in a study of 590 children taken from different samples with a mean age of 54.42 months ($SD = 13.57$), primarily white, and of middle and upper socioeconomic status (SES). Internal consistency (alpha) ranged from .55 to .80 when mothers were the respondents and from .40 to .74 when the fathers answered the questionnaires. A second study, also by Putnam & Rothbart (2006), was carried out with three more diverse samples as regards ethnicity and SES, and using data from applying the CBQ-SF itself. Sample 1 comprised of 138 primarily white children (mean age 68.0 months, $SD = 20.67$) with alpha values ranging from .54 to .79; Sample 2 included 289 children aged 3 of diverse ethnicity with alpha values ranging from .43 to .82; and Sample 3 was made up of 169 predominantly white children (mean age 50.17 months, $SD = 4.83$) with alpha values ranging from .46 to .82. Several short form scale scores showed unacceptable internal consistency when applied to diverse samples, but the coefficients improved when the questionnaires were administered to homogeneous samples. In any case, all the short form scale scores showed consistency across age and informant.

In addition, the 36-item very short form (CBQ-VSF) (Putnam & Rothbart, 2006) was designed to provide three orthogonal scale scores reflecting the three broader factors, with 12 items per broad dimension. The items retained were those exhibiting strong correlations with their associated factor and weak correlations with the other two. This CBQ-VSF has only been studied in a Dutch sample aged 6 to 8 (Sleddens et al., 2011) using data extracted from the Standard CBQ and, more recently (Sleddens et al., 2012) in 237 children aged 3 to 5, with mean age of 3.9 years ($SD = 0.8$), medium and high SES, and varied ethnicity. These studies found three broad dimensions with low or null inter-correlations and moderate internal consistency ranging from between .62 for Effortful Control and .78 for the same factor when using data extracted from the short form (Putnam & Rothbart, 2006), and between .74 and .78 on applying the CBQ-VSF (Sleddens et al., 2012).

The objective of this study is to evaluate the psychometric properties of both the CBQ-SF and CBQ-VSF scale scores in an extensive Spanish community sample representative of the general population of 3-year-old preschool children. The specific aims are to study the internal structure and the relationship with external variables in order to provide evidence of the reliability and validity of their scores for the assessment of young children's temperament in relation to psychopathology.

Method

Participants

The sample is part of the first assessment of a longitudinal research of vulnerability to psychopathology, specifically behavioral problems in preschool children. Based on the vulnerability/predisposition model, the study focuses on many biological and environmental risk factors for developing psychopathology, including temperament. The initial sample consisted of 2,283 children randomly selected from all registered preschoolers (age 3) in Barcelona for academic year 2009-10. The research was launched in a two-phase design

(Ezpeleta, de la Osa, & Domènech, in press). The proportion of families that agreed to participate in the first phase of screening was 58.7% ($N = 1,341$ families), and no differences were found by sex ($p = .95$) when comparing participants and refusals. However, the proportion of refusals was statistically higher for families in low socio-economic groups ($p < .001$). The screening for including children in the second phase was carried out with the parents' version of the Strengths and Difficulties Questionnaire for 3- and 4-year-olds, (SDQ³⁻⁴) (Goodman, 1997). In the second phase of the research, all the families of children with a positive screening score for behavioral problems ($n = 522$) and a random sample of 30% of children with a negative screening score ($n = 235$) were invited to continue. The final second-phase sample included 82.2% of the families invited to continue ($N = 622$ children); no statistical differences were found by sex ($p = .820$) or type of school ($p = .850$) when comparing participants and refusals in this second phase. The children's mean age was 2.97 ($SD = 0.16$); 310 were boys (49.8%) and 558 were white (88.9%). Table 1 shows the sociodemographic features of the sample and the (weighted) prevalence for DSM-IV disorders. Children with an intellectual disability or pervasive developmental disorders were excluded.

Measures

The *Child Behavior Questionnaire - Short Form* (CBQ-SF) and *Very Short Form* - (CBQ-VSF) (Putnam & Rothbart, 2006). The 94 items on the CBQ-SF are rated on a Likert-type scale of 1 (*extremely untrue of your child*) to 7 (*extremely true of your child*), plus a *not applicable* response option when the parents have not observed the child in the situation described. The items measure 15 theoretical dimensions of child temperament: Activity Level, Anger/Frustration, Approach/Positive Anticipation, Attentional Focusing, Discomfort, Falling Reactivity/Soothability, Fear, High Intensity Pleasure, Impulsivity, Inhibitory

Control, Lack shyness, Low Intensity Pleasure, Perceptual Sensitivity, Sadness, and Smiling and Laughter. As in Putnam and Rothbart (2006), the CBQ-VSF comprises 36 items extracted from the CBQ-SF, capturing the three broad dimensions mentioned earlier, namely NA, E/S, and EC. For both forms, scale scores are derived from the average of the corresponding items, after reversal where necessary. The CBQ-SF was available for 612 participants (98.4% of the sample). The respondents were parents (62.7% mothers, 8.0% fathers, and 29.3% mother-father pairs). The number of questionnaires answered just by fathers (8%) did not allow the data to be analyzed separately. Parents could choose between the Spanish or Catalan version of the instrument depending on the language they speak with each other. Both languages are indistinctly spoken by children at school and daily life and they are similar, as both are Romance and official languages in this region. The Spanish official version, already published on the authors' website, was used with their permission. We also obtained permission to adapt the questionnaire into Catalan, as we performed the study in a bilingual context. Two bilingual clinical psychologists translated the instrument following the widely accepted guidelines issued by the International Test Commission (2010). Experts in clinical psychology, also bilingual, checked the resulting version.

The *Diagnostic Interview of Children and Adolescents for Parents of Preschool Children and Young Children (DICA-PPYC)* (Ezpeleta, de la Osa, Granero, Domènech, & Reich, 2011). This test was used to assess children's psychopathology according to DSM-IV-TR taxonomy (American Psychiatric Association, 2000). The adaptation and validation for the Spanish preschool population showed sound psychometric properties (Ezpeleta et al., 2011). The diagnoses included in this study are presented in Table 1.

The *Strengths and Difficulties Questionnaire (SDQ³⁻⁴)*; Goodman, 1997). This questionnaire is for parents and teachers of 3- to 4-year-old children. It contains 22 items with the same content as the other versions, while two items about behavioral problems (lying and

stealing) were replaced by items reflecting oppositional behaviors (argumentativeness and spitefulness). Likewise, the wording for item 21 was modified (Can stop and think things out before acting). The items have three response choices (0: *not true*; 1: *somewhat true*; 2: *certainly true*). The official Spanish and Catalan versions of the SDQ³⁻⁴ were used (Ezpeleta, Granero, de la Osa, Penelo, & Domènech, 2013). This questionnaire was used to screen and select the sample. Table 5 includes the internal consistency of the current sample.

The *Children's Global Assessment Scale* (Shaffer et al., 1983) . The Spanish adaptation (Ezpeleta, Granero, & de la Osa, 1999) of this scale was used to assess global functional impairment based on children's psychopathology. Test-retest CCI for this measure was .65 and inter-rater reliability was > .70 for a sample aged 6 to 17. The total score was used, ranging from 0 (highest impairment) to 100 (lowest impairment).

The *Behavior Rating Inventory of Executive Function for Preschool Children* (BRIEF-P; (Gioia, Isquith, Guy, Kenworthy, & Baron, 2000). This test measures executive functions and was included in the general study due to the stated relationship between these variables and externalizing disorders, including ODD. Moreover, we included this measure in this study because a lack of self-regulation has been proposed as a possible precursor of antisocial behaviors in adulthood. It has 63 items and three response options (0: *never* to 2: *very often/always*) and is completed by teachers. For this study, we used the ISCI global index (the sum of the items included in inhibitory and emotional control scales), the FI index (the sum of shift and emotional control scales), the EMI index (the sum of working memory and plan-organize scales), and the GEC index (the sum for total scale). Table 5 includes internal consistency for the scale scores obtained in the sample used for this study (Ezpeleta, Granero, Penelo, de la Osa, & Doménech, in press)

Procedure

The project was approved by the ethics review committee of the authors' institution. The head teachers of the participating schools, as well as the children's parents, received a complete description of the study. The families were recruited at the schools and gave written consent. All parents of children in grade P3 (aged 3) in the participating schools were invited to answer the SDQ³⁻⁴ at home and return it to the schools. The families who agreed and met the screening criteria were contacted by telephone and interviewed at the school. The interviewers were previously trained and were blind to the children's screening group. Immediately after the interview, the interviewer completed the CGAS, the parents answered the CBQ-SF, and the teachers answered the BRIEF-P questionnaire.

The diagnostic interview was administered twice, at ages 3 and 4. The attrition rate between the two assessments was 2.9% (18 participants lost). Children from each school were assessed about 12 months after the first assessment.

Statistical Analysis

The statistical analysis was carried out with SPSS19 and Mplus6. Because of the multistage design (the sample analyzed in this work was selected through a screening procedure), analyses in SPSS were carried out into the Complex Samples system (creating a plan file with sampling weights inversely proportional to the possibility of participant selection in the second phase of the study) and analyses in Mplus6 program were weighted using the same algorithm.

The 15 CBQ scale scores of the CBQ-SF were subjected to Exploratory Structural Equation Modeling (ESEM; Asparouhov & Muthén, 2009) with Mplus6, using the robust maximum likelihood estimator (MLR). An exploratory approach was adopted and justified by the absence of previous researches which reported the internal structure for the CBQ-SF and because this approach permits the initial estimation of the factor loadings of all items in all

the factors (whereas CFA implies the additional problem of fixing the possible cross-loadings to 0 (Marsh et al., 2010; Marsh, Nagengast, & Morin, 2013). Goodness of fit was assessed with the common fit indices (Jackson, Gillaspay, & Purc-Stephenson, 2009): χ^2 , Tucker and Lewis Index (TLI), Comparative Fit Index (CFI), and Root Mean Square Error of Approximation (RMSEA). The following thresholds were considered good fit to: RMSEA < .06, TLI and CFI > .95. Moderate and reasonable fit was considered for, RMSEA < .08 and TLI and CFI > .90 (Marsh, Hau, & Wen, 2004).

In addition to assess the internal structure of the CBQ-SF, we set out to examine the internal structure of the CBQ-VSF by subjecting the 36 CBQ items to: a) confirmatory factor analysis with MPlus6 (with the aim to assess the two models for CBQ-VSF previously published in older children) (Putnam & Rothbart, 2006; Sleddens et al., 2011); and b) principal axis factor analysis with oblimin oblique rotation, as in Sleddens et al. (2011) with SPSS19. For the latter, and to deal with the missing values (0.47%), we used the expectation-maximization (EM) method of data imputation. Once the internal structure of the CBQ was determined, the internal consistency of the derived scores was measured with Cronbach's alpha. In addition, correlations between analogous broad dimensions derived from the SF and VSF were calculated to assess the equivalence between the two forms. The similarity between our results on the structure of CBQ-VSF and those published by Sleddens et al. (2011), both using principal axis factor, was examined. Since the authors did not provide factor loadings below .25, it was impossible to calculate Tucker's *c* coefficient of congruence. Therefore, we calculated Cohen's Kappa statistic (κ ; Cohen, 1960) as proposed by Guadagnoli and Velicer (1991), where salient loadings were defined as above .30 without regard to sign, and low non-reported factor loadings were substituted by zero values. For analogous factor pairs, $\kappa > .75$ represents excellent agreement, values between .40 and .75 are

indicators of fair to good agreement, and $\kappa < .40$ represents poor agreement; for non-analogous factor pairs, $\kappa < 0$ is expected (Guadagnoli & Velicer, 1991).

The association between CBQ dimensions and SDQ³⁻⁴, and BRIEF-P was calculated with Pearson's correlations (r). Due to the large sample size and the high statistical power, low correlation values tended to be statistically significant, so only r -coefficients with good effect sizes ($|r| \geq .30$) were considered relevant.

The discriminative capacity of the CBQ dimensions (screening tool) for identifying children who presented specific DSM disorders (measured through diagnostic interview, considered the reference measure or diagnostic measure) was estimated as an additional measure of the convergent validity of the screening instrument. The DSM disorders analyzed were disruptive disorders (ADHD, ODD, and CD) and anxiety disorders (separation disorder, generalized disorder, phobias, and somatization). Binary logistic regressions were used (CBQ dimension scores were entered as the independent variables and the presence-absence of disorders were considered the dependent variables –criteria-), adjusted for the covariates sex and comorbidities other than those included in the model. The three broad CBQ dimensions were entered in the regressions together in order to measure the specific contribution of each factor to the possibility of each disorder appearing. The area under the receiver operator curve (AUC) measured the discriminative accuracy of the models (AUCs below .65 were considered poor, between .65 and .70 moderate, and above .70 good). This procedure was also used to measure the predictive accuracy of the CBQ dimensions and the presence of impairment (CGAS scores below 70).

Results

Internal Structure and Internal Consistency of the CBQ-SF Scores

The goodness-of-fit indices of the 3-factor model of the CBQ-SF were satisfactory: $\chi^2(63) = 165.4$, CFI = .94, TLI = .90, RMSEA = .052 (90% confidence interval [CI]: .042;

.061). The factor loadings and factor correlations are shown in Table 2 (left). The Negative Affect (NA) factor was defined by loadings for the scales of anger/frustration, discomfort, fear, sadness, shyness. Soothability/falling reactivity loaded negatively in the NA factor. The Surgency/Extraversion (SE) factor was defined by loadings for the scales of activity level, impulsivity, high intensity pleasure, and approach/positive anticipation; and the Effortful Control (EC) factor was defined primarily by loadings for the scales of attentional focusing, inhibitory control, low intensity pleasure, perceptual sensitivity, and smiling and laughter. Therefore, all the scales loaded significantly and principally in the factor with which they were associated, as in Sleddens et al. (2011).

Table 2 (right), also presents results for internal consistency scores derived from the three broad dimension's. Internal consistency for the three broad dimensions ($\alpha \geq .79$) was satisfactory (Nunnally & Bernstein, 1994). For the 15 primary scale scores, internal consistency ranged from .48 (Approach/Positive Anticipation) to .79 (Shyness). All mean inter-item correlations were above .15, except for Sadness (.13) and Approach/Positive Anticipation (.14).

Internal Structure and Internal Consistency of the CBQ-VSF Scores

Using confirmatory factor analysis, we failed to replicate results for CBQ-VSF published previously (Putnam & Rothbart, 2006; Sleddens et al., 2011), since the factor loadings and fit indexes did not reach adequate solutions (data not shown; available upon request). In contrast, principal axis factor analysis for the CBQ-VSF yielded 12 factors with eigenvalues greater than one. After forcing a three-factor solution, as suggested by the scree plot test, a more interpretable and plausible factor structure was found representing the three broad dimensions expected. Table 3 presents the factor loadings (pattern matrix) and internal consistency ($\alpha \geq .65$). For the 36 items factorized, three of EC, three of NA, and four of S/E, showed crossloadings and/or low coefficients in the expected factor. Direct scale scores of

the 36-item CBQ-VSF were strongly correlated with the corresponding direct scale scores of the CBQ-SF: r between .83 for EC and .93 for NA ($p < .001$). Factor correlations were null.

Regarding the similarity of the factor structure of the CBQ-VSF with the solution published by Sleddens et al. (2011) using exactly the same technique, there was excellent agreement for the NA scale ($\kappa = .83$) and satisfactory agreement for the EC ($\kappa = .74$) and SE ($\kappa = .56$) scales. Values for non-analogous factor pairs were also as expected ($\kappa \leq -.06$).

Distribution of Raw Scores

All the results from now on are based on direct scale scores derived from the average of the corresponding items, after reversal where necessary. Table 4 shows the weighted means and standard deviations for the CBQ dimensions and effect size of mean differences across sex based on Cohen's d coefficient. These results show that all differences by sex were within the range of small effect sizes ($|d|$ below 0.50).

Association between CBQ Scale Scores and External Variables

Table 5 shows Pearson's correlations between the CBQ dimension scores and SDQ³⁻⁴ and BRIEF-P scores. Comparing CBQ-SF and SDQ³⁴ measures, CBQ-Negative Affect scale scores were positively correlated with SDQ-parents Emotional and Internalizing scale scores; CBQ-Surgency Extraversion dimension scores correlated with SDQ-parents ADHD, Externalizing and Total scores; and CBQ-Effortful Control scale negatively correlated with SDQ-parents ADHD, Externalizing, and Total scale scores, and with SDQ-teachers ADHD measures. The comparison of CBQ-VSF and SDQ³⁴ showed that CBQ-Negative Affect scores positively correlated with SDQ-parents Emotional and Internalizing scores; CBQ-Surgency scores were positively associated with SDQ-parents Externalizing scores.

Compared with the BRIEF-P scores, CBQ-SF Effortful Control scores negatively correlated with EMI and GEC scores, while CBQ-VSF Effortful was negatively associated with EMI scores.

Predictive Accuracy of CBQ Dimensions and DSM-IV Disorders and Impairment

Table 6 (left) shows logistic regressions adjusted for the covariates sex and comorbidities other than those included in the model, indicating the discriminative accuracy of the CBQ dimensions for the presence of DSM-IV disorders. All the models achieved significant discriminative accuracy, with AUC values ranging between moderate (around .65 for any DSM-IV disorder) and excellent (above .80 for ADHD, social phobia, and separation anxiety at age 4 predicted by the CBQ-VSF). The main cross-sectional results in Table 6 indicate that: a) the Negative Affect dimension was associated with any disorder, oppositional defiant disorder, anxieties, and impairment; b) CBQ-Surgency/Extraversion was a significant predictor for any disorder, disruptive disorders, and social phobia, and c) the CBQ-Effortful dimension significantly contributed to disruptive disorders and impairment, as well as to specific phobia at age 4. The data also show that all the significant OR coefficients for the contribution of the CBQ-Negative and CBQ-Surgency/Extraversion dimensions were above one, indicating that high values in these dimensions increase the probability for the presence of the disorders; the only exception was Surgency/Extraversion in the model for social phobia, whose OR values were significant and lower than one (higher scores in this scale were predictive of the absence of this disorder). In contrast, the CBQ-Effortful dimension yielded significant OR coefficients lower than one (the higher the dimension score, the lower the risk of the disorder).

The results of the right column of Table 6 show the predictive validity of the CBQ dimensions assessed at age 3 to detect the presence of DSM-IV disorders at age 4 (results adjusted for comorbidities other than those included in the model, and for the presence of the specific disorder at age 3). For both formats (short and very short), the CBQ-Negative affect scale was predictive of the presence of any disorder, disruptive disorder (specifically ODD), any anxiety disorder, phobias, and impairment. CBQ-Surgency measured with the short

format was predictive of disruptive disorder (ADHD), separation anxiety, and phobias, and measured with the very short format it was predictive of phobias. CBQ-Effortful was only predictive of disorders at age 4 measured with the short form, specifically for ADHD, any anxiety disorder and specific phobia.

Discussion

This is the first study exploring the psychometric properties of the Spanish CBQ-SF and CBQ-VSF scores in 3-year-old preschool children from a general population with a longitudinal design. It is also the first study with an extensive and homogeneous sample as regards age (3 years old). This is the limit age to apply this instrument, and age has proved to be an important aspect in the items' appropriateness. In the future, the longitudinal design of the study that data belong to will permit us to study the stability of the measure and the developmental pathways of the relationship between temperament and psychopathology. The results showed that for both forms, the three-broad dimension structure was replicated and the scale scores derived presented acceptable validity and reliability for assessing temperament in preschoolers.

We wanted to explore the utility of the CBQ-SF scores for examining specific temperament characteristics as well as obtaining information about the three broad factors that have been consistently found in the standard CBQ. The potential utility of examining personality, in terms of both higher-order factors and lower-order facets, has been studied in the sense that different components of the same domain may be related to psychopathology (De Clerck, De Fruyt, Van Leeuwen, & Mervielde, 2006; Salekin, Debus, & Baker, 2010). The CBQ-SF offers this possibility more efficiently than the standard form, which is more time-consuming. No other study has done this, so we shall discuss the internal structure and internal consistency of the CBQ-SF scores and compare them with those obtained by Rothbart et al. (2001) in a sample of 149 children aged 3 and 517 children aged 4-5 using the

standard CBQ. The factors found in this large Spanish sample are similar to those found previously under the consensus-based labels of NA, SR, and EC. For the CBQ-SF, we replicated the structure obtained by Rothbart et al. (2001) using the standard CBQ. A first factor dovetailed with the global dimension of NA. Soothability did not load in the EC factor as it does in older samples, while shyness behaved the same as in older samples and in the Dutch validation of the CBQ-VSF by Sleddens et al. (2011) with a sample of children aged 6 to 8. A second factor was identifiable as the S/E dimension by loading on exactly the same scales as in the standard form. The third factor corresponding to the global EC dimension matches what was obtained with the original standard form (Rothbart et al., 2001). Congruence coefficients indicate the high degree of agreement between the CBQ-VSF and the Dutch internal structure obtained by Sleddens et al. (2011). Considering the difference in the number of items (more than 100) between the standard CBQ and the CBQ-SF and the strong similarity in structure, along with the internal consistency of the derived dimension scores, these results are quite encouraging in terms of using the CBQ-SF in temperament research.

As regards correlations between the three broad factor scores of CBQ-SF, EC correlated weakly and negatively with SE but was uncorrelated with NA, as in the Chinese sample with children aged 6 to 7 (Ahadi et al., 1993). Cultural differences determining which types of behavior should be inhibited may explain these differences (Rothbart et al., 2001), although the magnitudes of the differences were low.

Internal consistency of the three broad scores of the CBQ-VSF was similar for S/E, but lower for NA and EC than in the results reported in the original study. Sleddens et al. (2011), also working with data from the standard CBQ, obtained lower values for NA (.72) but also higher alpha values for SE (.76) and EC (.72). Again, their sample was older than ours, and the two weakest scales in our study are EC and SE, probably due to the younger and

narrower age range of our sample, leading to lower variability in scores and hence to a drop in the coefficient values.

We compared our reliability results for the CBQ-SF with those from one of the three samples used by Putnam & Rothbart, (2006) ($N = 289$ 3-year-olds), similar to ours in age but with different ethnicity (48% African-American) and socio-economic status (49% low SES) profiles. In the original study, alpha values ranged from .43 (sadness) to .82 (shyness). Four scale scores in our study (approach/positive participation, sadness, low intensity pleasure, and smiling/laughter) failed to yield acceptable internal consistency. This is in line with the original study.

As regards the relationship between scores derived from both instruments and external variables, it was mainly as expected and as reported in the literature on the relationship between temperament and functioning and psychopathology. The convergence of CBQ dimension scores and SDQ scores is consistent with other findings (Putnam & Rothbart, 2006) that report high levels of externalizing behavior in highly surgent toddlers (Putnam & Stifter, 2005) and a relationship between lack of self-control and behavior problems (Hill, Degnan, Calkins, & Keane, 2006; Kochanska & Knaack, 2003; Putnam et al., 2008).

The link between executive attention and effortful control has largely been validated (Putnam & Stifter, 2008). The convergence between EC scores and SDQ measures as regards attention and externalizing problems also illustrates this link and provides further support for the CBQ-SF as a useful questionnaire for the early detection of risk factors for developmental problems. CBQ scores also correlated for the dimensions expected (EC) with scores from a specific questionnaire on executive functions, such as the BRIEF-P. NA temperament characteristics are related to any kind of mental problem. The relationship between inhibited temperament and internalizing problems and anxiety disorders has consistently been reported in the literature (Lemery, Essex, & Smider, 2002; Putnam & Stifter, 2005).

The association between CBQ dimensions in any of its short versions and theoretically-related DSM-IV disorders also constitutes evidence of the validity of the measure for use in studies about preschoolers' developmental psychopathology. The association between the NA dimension and ODD (usually considered a conduct problem) is noteworthy and reflects the multiple dimensions of this disorder (Burke, Hipwell, & Loeber, 2010; Ezpeleta et al., 2012). The predictive value of certain temperament difficulties at age 3 for problems at age 4, even when adjusted for the presence of disorders at age 3, gives those difficulties potential clinical utility, and this finding is in line with the idea of continuity in temperament and its link with pathology at very young ages (Asendorpf et al., 2008; Cipriano & Stifter, 2010). This predictive capacity applies to both disruptive and anxiety disorders, and to daily functioning. The NA scores of CBQ-VSF in particular have great predictive capacity.

Leerkes, Paradise, O'Brien, Calkins, & Lange (2008) found that emotional understanding and control were more closely related to social and academic functioning than cognitive abilities are. The clear association between impairment measures and all of the scales (except Surgency) of both questionnaire forms reinforces the idea that early temperament problems are a clear risk factor for proper functioning (Healey et al., 2010).

The study does have some limitations. Our results for the CBQ-VSF were taken from applying the CBQ-SF, and this could affect the response process, potentially leading to a drop in the reliability and validity of CBQ-VSF scores, as pointed out by Sleddens et al. (2012) in their discussion of the disadvantage of using such extracted data. Likewise, the fact that participation in the study was voluntary meant that the SES of the sample was higher than that of the general population, since few families with low socioeconomic status participated, and this could have led to bias. The participants' age made difficult to find instruments validated for the preschooler age range, as was the case with the functioning measure used, the CGAS, which is validated in Spanish population beyond the age 6. To the best of our

knowledge, this is the largest sample ever used in a longitudinal study on data from the 3-year-old general population using the CBQ-SF and CBQ-VSF in a longitudinal design. Scores obtained with the short and very short forms of the CBQ appear to show that they have acceptable psychometric properties at an early age. CBQ -VSF is useful for screening and detecting the more salient aspects of temperament and those related to psychopathology; moreover, they are efficient, leading to savings in time and economic resources while still providing information on lower-order aspects and thus allowing for research about the structure of child personality and its relationship to psychopathology, as in the case of CBQ-SF. The factor structure of temperament in children is quite stable across samples from different countries, though more studies should be carried out to improve our understanding of certain differences in certain factor loadings between cultures, informants, and ages.

Acknowledgements

Funding was from the Spanish Ministry of Science and Innovation grants PSI2009-07542 and PSI2012-32695 from the Ministry of Economy and Competitiveness. We would like to thank the participating schools and families.

References

- Ahadi, S. A., Rothbart, M. K., & Ye, R. M. (1993). Childrens temperament in the us and china - similarities and differences. *European Journal of Personality, 7*(5), 359-378. doi:10.1002/per.2410070506
- Aksan, N., Goldsmith, H. H., Smider, N. A., Essex, M. J., Clark, R., Hyde, J. S., . . . Vandell, D. L. (1999). Derivation and prediction of temperamental types among preschoolers. *Developmental Psychology, 35*(4), 958-971. doi:10.1037//0012-1649.35.4.958
- American Psychiatric Association. (2000). *DSM-IV Diagnostic and statistical manual of mental disorders* (4th Text Revised ed.). Washington, DC: American Psychiatric Press
- Asendorpf, J. B., Denissen, J. J. A., & van Aken, M. A. G. (2008). Inhibited and aggressive preschool children at 23 years of age: Personality and social transitions into adulthood. *Developmental Psychology, 44*(4), 997-1011. doi:10.1037/0012-1649.44.4.997
- Asparouhov, T. & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling, 16*, 397-438. doi:10.1080/10705510903008204
- Blair, C. (2002). School readiness - integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist, 57*(2), 111-127. doi:10.1037//0003-066X.57.2.111
- Burke, J. D., Hipwell, A. E., & Loeber, R. (2010). Dimensions of oppositional defiant disorder as predictors of depression and conduct disorder in preadolescent girls. *Journal of the American Academy of Child and Adolescent Psychiatry, 49*(5), 484-492. doi:10.1016/j.jaac.2010.01.016
- Caspi, A., Harrington, H., Milne, B., Amell, J. W., Theodore, R. F., & Moffitt, T. E. (2003). Children's behavioral styles at age 3 are linked to their adult personality traits at age 26. *Journal of Personality, 71*(4), 495-513.
- Cipriano, E. A., & Stifter, C. A. (2010). Predicting preschool effortful control from toddler

- temperament and parenting behavior. *Journal of Applied Developmental Psychology*, 31(3), 221-230. doi:10.1016/j.appdev.2010.02.004
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
- De Clercq, B., De Fruyt, F., Van leeuwen, K. & Mervielde, I. (2006). The structure of maladaptative personality traits in childhood: a step towards an integrative developmental perspective for DSM-V. *Journal of Abnormal Psychology*, 115, 639-657. doi:10.1037/0021-843X.115.4.639
- Egger, H. L., & Angold, A. (2006). Common emotional and behavioral disorders in preschool children: Presentation, nosology, and epidemiology. *Journal of Child Psychology and Psychiatry*, 47(3), 313-337.
- Eggers, K., De Nil, L. F., & Van den Bergh, B. R. H. (2009). Factorial temperament structure in stuttering, voice-disordered, and typically developing children. *Journal of Speech Language and Hearing Research*, 52(6), 1610-1622. doi:10.1044/1092-4388(2009/07-0065)
- Ezpeleta, L., Granero, R., & de la Osa, N. (1999). Evaluación del deterioro en niños y adolescentes a través de la Children's Global Assessment Scale (CGAS). *Revista de Psiquiatría Infanto-Juvenil*, 1, 18-26.
- Ezpeleta, L., Granero, R., de la Osa, N., Penelo, E., & Domènech, J. M. (2013). Psychometric properties of the Strengths and Difficulties Questionnaire³⁻⁴ in 3-year-old preschoolers. *Comprehensive Psychiatry*, 54, 282-291.
doi:[10.1016/j.comppsy.2012.07.009](https://doi.org/10.1016/j.comppsy.2012.07.009)
- Ezpeleta, L., Granero, R., Penelo, E., de la Osa, N., & Domènech, J. M. (in press). Behavior Rating Inventory of Executive Functioning–Preschool (BRIEF-P) applied to teachers: Psychometric properties and usefulness for disruptive disorders in 3-year-old

- preschoolers. *Journal of Attention Disorders*. doi:10.1177/1087054712466439
- Ezpeleta, L., de la Osa, N., Granero, R., Domènech, J. M., & Reich, W. (2011). The diagnostic interview of children and adolescents for parents of preschool and young children: Psychometric properties in the general population. *Psychiatry Research*, *190*, 137-144. doi:10.1016/j.psychres.2011.04.034
- Ezpeleta, L., de la Osa, N., & Domènech, J.M. (in press). Prevalence of DSM-IV disorders, comorbidity and impairment in 3-year-old Spanish preschoolers. *Social Psychiatry*. doi:10.1007/s00127-013-0683-1.
- Gartstein, M. A., & Rothbart, M. (2003). Studying infant temperament via the Revised Infant Behavior Questionnaire. *Infant Behavior & Development* *26*, 64-86. doi: 10.1016/S0163-6383(02)00169-8
- Gioia, G. A., Isquith, P. K., Guy, S. C., Kenworthy, L., & Baron, I. S. (2000). Test review: Behavior rating inventory of executive function. *Child Neuropsychology*, *6*(3), 235-238. doi:10.1076/chin.6.3.235.3152
- Goldsmith, H. H., Buss, K. A., & Lemery, K. S. (1997). Toddler and childhood temperament: Expanded content, stronger genetic evidence, new evidence for the importance of environment. *Developmental Psychology*, *33*(6), 891-905. doi:10.1037//0012-1649.33.6.891
- Goodman, R. (1997). The strengths and difficulties questionnaire. *Journal of Child Psychology and Psychiatry*, *38*, 581-586.
- Guadagnoli, E. & Velicer, W. (1991). A comparison of pattern matching indices. *Multivariate Behavioral Research*, *26*, 323-343.
- Healey, D. M., Brodzinsky, L. K., Bernstein, M., Rabinovitz, B., & Halperin, J. M. (2010). Moderating effects of neurocognitive abilities on the relationship between temperament

and global functioning. *Child Neuropsychology*, *16*(1), 20-31.

doi:10.1080/09297040902984490

Hill, A. L., Degnan, K. A., Calkins, S. D., & Keane, S. P. (2006). Profiles of externalizing behavior problems for boys and girls across preschool: The roles of emotion regulation and inattention. *Developmental Psychology*, *42*(5), 913-928. doi:10.1037/0012-1649.42.5.913

International Test Commission (2010). *International Test Commission Guidelines for Translating and Adapting Tests*. [<http://www.intestcom.org>].

Jackson, D. L., Gillaspay, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods*, *14*, 6-23. doi:10.1037/a0014694

Kochanska, G., Devet, K., Goldman, M., Murray, K., & Putnam, S. P. (1994). Maternal reports of conscience development and temperament in young-children. *Child Development*, *65*(3), 852-868. doi:10.1111/j.1467-8624.1994.tb00788.x

Kochanska, G., & Knaack, A. (2003). Effortful control as a personality characteristic of young children: Antecedents, correlates, and consequences. *Journal of Personality*, *71*(6), 1087-1112. doi:10.1111/1467-6494.7106008

Leerkes, E. M., Paradise, M., O'Brien, M., Calkins, S. D., & Lange, G. (2008). Emotion and cognition processes in preschool children. *Merrill-Palmer Quarterly-Journal of Developmental Psychology*, *54*(1), 102-124. doi:10.1353/mpq.2008.0009

Lemery, K. S., Essex, M. J., & Smider, N. A. (2002). Revealing the relation between temperament and behavior problem symptoms by eliminating measurement confounding: Expert ratings and factor analyses. *Child Development*, *73*(3), 867-882. doi:10.1111/1467-8624.00444

- Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*, 320-341. doi:10.1207/s15328007sem1103_2
- Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment, 22*, 471-491. doi:10.1037/a0019227
- Marsh, H. W., Nagengast, B., & Morin, A. J. S. (2013). Measurement invariance of Big-Five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and La Dolce Vita effects. *Developmental Psychology, 49*, 1194-1218. doi:10.1037/a0026913
- Muris, P., & Ollendick, T. H. (2005). The role of temperament in the etiology of child psychopathology. *Clinical Child and Family Psychology Review, 8*(4), 271-289. doi:10.1007/s10567-005-8809-y
- Nunnally, J. C. & Bernstein, I. J. (1994). *Psychometric theory* (3rd. ed.). New York: McGraw-Hill.
- Nygaard, E., Smith, L., & Torgersen, A. M. (2002). Temperament in children with down syndrome and in prematurely born children. *Scandinavian Journal of Psychology, 43*(1), 61-71. doi:10.1111/1467-9450.00269
- Putnam, S. P., & Stifter, C. A. (2005). Behavioral approach-inhibition in toddlers: Prediction from infancy, positive and negative affective components, and relations with behavior problems. *Child Development, 76*(1), 212-226. doi:10.1111/j.1467-8624.2005.00840.x
- Putnam, S. P., & Rothbart, M. K. (2006). Development of short and very short forms of the children's behavior questionnaire. *Journal of Personality Assessment, 87*(1), 102-112. doi:10.1207/s15327752jpa8701_09

- Putnam, S.P., Gartstein, M. A., & Rothbart, M.K. (2006). Measurement of fine-grained aspects of toddler temperament: The Early Childhood Behavior Questionnaire. *Infant Behavior & Development* 29, 386-401. doi:10.1016/j.infbeh.2006.01.004
- Putnam, S. P., Rothbart, M. K., & Gartstein, M. A. (2008). Homotypic and heterotypic continuity of fine-grained temperament during infancy, toddlerhood, and early childhood. *Infant and Child Development*, 17(4), 387-405. doi:10.1002/ICD.582
- Putnam, S. P., & Stifter, C. A. (2008). Reactivity and regulation: The impact of Mary Rothbart on the study of temperament. *Infant and Child Development*, 17(4), 311-320. doi:10.1002/icd.583
- Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The children's behavior questionnaire. *Child Development*, 72(5), 1394-1408. doi:10.1111/1467-8624.00355
- Rothbart, M.K., & Derryberry, D. (1981). Development of individual differences in temperament. In M.E. Lamb & A. L. Brown (Eds.), *Advances in developmental psychology* (Vol.1, pp.37-86). Hillsdale, NJ: Erlbaum.
- Salekin, R.T., Debus, S.A., & Baker, E.D. (2010). Adolescents psychopathy and the five factor models: domain and facet analysis. *Journal of Psychopathology and Behavioral Assessment*, 32, 4, 501-514. doi:10.1007/s10862-010-9192-7
- Shaffer, D., Gould, M. S., Brasic, J., Ambrosini, P., Fisher, P., & Bird, H. (1983). A children's global assessment scale (CGAS). *Archives of General Psychiatry*, 40, 1228-1231.
- Sleddens, E. F. C., Kremers, S. P. J., Candel, M. J. J. M., De Vries, N. N. K., & Thijs, C. (2011). Validating the children's behavior questionnaire in Dutch children: Psychometric properties and a cross-cultural comparison of factor structures. *Psychological Assessment*, 23(2), 417-426. doi:10.1037/a0022111

- Sleddens, E. F. C., Hughes, S. O., O'Connor, T. M., Beltran, A., Baranowski, J.C., Nicklas, T. A., & Baranowski, T. (2012). The Children's Behavior Questionnaire very short scale: Psychometric properties and development of a one-item temperament scale. *Psychological Reports, 110*, 197-217.
- Tackett, J.L. (2006). Evaluating models of the personality-psychopathology relationship in children and adolescents. *Clinical Psychological Review, 26*, 584-599.
doi:10.1016/j.cpr.2006.04.003.
- Tackett, J.L., Balsis, S., Oltmann, T.F., & Krueger, R.F. (2009). A unifying perspective on personality pathology across lifespan: Developmental considerations for DSM-V. *Development and Psychopathology, 21*, 687-713. doi:10.1017/S095457940900039X
- Wolfe, C. D., & Bell, M. A. (2007). The integration of cognition and emotion during infancy and early childhood: Regulatory processes associated with the development of working memory. *Brain and Cognition, 65*(1), 3-13. doi:10.1016/j.bandc.2006.01.009

Table 1.

Sociodemographic and clinical data for the sample (N = 622).

Child age (mean; <i>SD</i>)		2.97 (0.16)
Child sex (n; %)	Male	310 (49.8%)
Child race/ethnicity (n; %)	White	553 (88.9%)
	American Hispanic	49 (7.9%)
	African	2 (0.3%)
	Asian	6 (1.0%)
	Other	12 (1.9%)
Mother's age (mean; <i>SD</i>)		36.4 (4.7)
Father's age (mean; <i>SD</i>)		38.6 (5.8)
Family socioeconomic status	High	205 (33.0%)
(Hollingshead, 1975)	Mean-high	195 (31.4%)
	Mean	88 (14.1%)
	Mean-low	99 (15.9%)
	Low	35 (5.6%)
DSM-IV disorders and impairment		<i>N</i> (weighted %)
Any disorder		245 (34.4%)
Any disruptive disorder		87 (10.0%)
ADHD		33 (3.6%)
Oppositional defiant disorder		61 (6.9%)
Conduct disorder		10 (1.4%)
Mood disorders		4 (0.4%)
Any anxiety disorder		59 (7.7%)
Separation anxiety		18 (2.1%)
Generalized anxiety		1 (0.1%)
Specific phobia		27 (3.6%)
Social phobia		16 (1.9%)
Impairment: CGAS < 70		107 (14.5%)

Note. *SD*: standard deviation.

Table 2.

Results for the 15-scale and 3-factor SF-CBQ models with ESEM and internal consistency.

Scales	Factor	Factor loadings			Internal consistency	
		Negative affect	Surgency/extraversion	Effortful control	No. items	Cronbach's α^*
Negative affect					37	.85
Anger/Frustration		.609	.236	-.222	6	.73
Discomfort		.534	-.013	.127	6	.68
Fear		.558	-.047	-.044	6	.64
Soothability/Falling reactivity		-.462	.004	.322	6	.63
Sadness		.660	.211	.047	7	.51
Shyness		.471	-.231	-.090	6	.79
Surgency/extraversion					26	.83
Activity level		-.001	.543	-.448	7	.74
Impulsivity		-.331	.658	-.020	6	.61
High intensity pleasure		-.054	.578	-.174	6	.65
Approach/Positive anticipation		.203	.614	.005	6	.48
Effortful control					32	.79
Attentional focusing		.053	-.074	.640	6	.69
Inhibitory control		-.048	-.197	.697	6	.62
Low intensity pleasure		.001	.400	.550	8	.55
Perceptual sensitivity		.132	.442	.473	6	.69
Smiling and laughter		-.152	.318	.346	6	.57
Factor correlations						
	Surgency/extraversion	-.03				
	Effortful control	.04	-.29			

Note. N=612. Bold: factor loadings $\geq .35$; shaded cells indicate the factor in which the scale was assigned, taken into account the content. *Based on shaded scales.

Table 3.

Results for the 36-item and 3-factor CBQ-VSF model with principal axis factor and oblimin rotation (pattern matrix), and internal consistency.

VSF Item No. and CBQ-SF/CBQ-Standard scale*	(N = 612)			
	KMO % explained variance after rotation	Negative affect	Surgency/ extraversion	Effortful control
		.76 22.6%		
02. Anger/Frustration [30/62]	.60	-.26	.01	
32. Anger/Frustration [40/78]	.49	-.12	-.07	
29. Discomfort [49/101] (reversed)	-.34	-.09	-.08	
08. Sadness [20/44]	.50	-.08	.03	
17. Sadness [27/55]	.43	-.04	.22	
05. Discomfort [29/61]	.47	-.01	.15	
14. Soothability/Falling reactivity [34/68]	.46	-.07	-.09	
23. Soothability/Falling reactivity [75/150]	.55	-.04	-.13	
11. Fear [17/40]	.39	-.08	-.02	
20. Discomfort [91/190] (reversed)	-.22	-.12	.00	
26. Fear [35/70] (reversed)	-.27	-.04	.03	
35. Sadness [31/64]	.43	-.08	.11	
07. Impulsivity [28/59]	-.02	-.59	.12	
19. Impulsivity [36/71] (reversed)	.46	.39	-.12	
31. Impulsivity [43/90] (reversed)	.26	.24	-.12	
34. Shyness [70/143] (opposite to item 10)	.37	.27	.00	
10. Shyness [11/23] (opposite to items 22-34)	-.29	-.28	.09	
22. Shyness [37/74] (opposite to item 10)	.46	.25	-.02	
01. Activity level [1/1]	.30	-.32	-.16	
13. Activity level [50/102] (reversed)	.04	.60	.21	
04. High intensity pleasure [4/8]	.13	-.49	.09	
25. Activity level [85/172]	.19	-.48	.01	
28. High intensity pleasure [78/159] (reversed)	.10	.40	.10	
16. High intensity pleasure [69/139]	-.03	-.26	.17	
18. Inhibitory control [67/136]	-.20	.26	.46	
30. Inhibitory control [73/147]	.00	.22	.11	
06. Inhibitory control [45/63]	.04	.04	.36	
03. Attentional focusing [62/125]	.03	.26	.45	
21. Low intensity pleasure [76/151]	-.19	-.07	.41	
12. Perceptual sensitivity [13/31]	.06	-.07	.51	
24. Perceptual sensitivity [47/98]	.05	-.12	.54	
15. Attentional focusing [71/144]	.00	.27	.43	
36. Perceptual sensitivity [32/65]	.09	-.06	.54	
09. Low intensity pleasure [72/146]	-.09	.02	.48	
27. Attentional focusing [89/186]	.20	-.03	.18	
33. Low intensity pleasure [94/164]	.09	-.22	.25	
Factor correlations				
Surgency/extraversion	.04			
Effortful control	.01	.01		
Internal consistency: α (no. items)**	.77 (15)	.65 (9)	.66 (12)	

Note. N=612. *SF/Standard form item numeration in square brackets.

**Based on shaded scales, which indicate the factor in which the scale was assigned, taken into account the content.

Table 4.

Distribution of CBQ raw scores: means (standard deviation) and comparison by sex.

	Total (<i>n</i> = 612)		Girls (<i>n</i> = 306)		Boys (<i>n</i> = 306)		Sex <i> d </i>
	Mean	<i>SD</i>	Mean	<i>SD</i>	Mean	<i>SD</i>	
CBQ-SF							
Negative affect	3.73	0.73	3.75	0.73	3.70	0.73	0.07
Anger/Frustration	4.24	1.16	4.14	1.16	4.34	1.16	0.17
Discomfort	4.20	1.16	4.40	1.08	3.99	1.20	0.36
Fears	3.59	1.25	3.64	1.27	3.54	1.23	0.08
Reactivity-Soothability	4.89	1.03	4.98	1.05	4.81	1.01	0.17
Sadness	3.80	0.91	3.79	0.92	3.80	0.90	0.01
Shyness	3.44	1.31	3.54	1.31	3.34	1.30	0.15
Surgency-extraversion	4.39	0.77	4.31	0.77	4.47	0.76	0.21
Activity level	4.01	1.15	3.83	1.17	4.20	1.10	0.33
Impulsivity	4.38	0.99	4.35	0.95	4.40	1.03	0.05
High intensity pleasure	4.42	1.11	4.28	1.10	4.56	1.11	0.25
Approach-Positive anticipation	4.75	0.77	4.76	0.80	4.73	0.74	0.04
Effortful control	5.32	0.58	5.41	0.55	5.24	0.60	0.30
Attentional Focusing	4.88	1.08	4.98	1.03	4.79	1.13	0.18
Inhibitory control	4.82	1.00	4.93	1.00	4.70	1.00	0.23
Low intensity pleasure	5.84	0.66	5.92	0.65	5.77	0.67	0.23
Perceptual sensitivity	5.50	1.00	5.66	0.93	5.34	1.04	0.32
Smiling and laughter	5.57	0.81	5.55	0.83	5.59	0.79	0.05
CBQ-VSF							
Negative affect	3.65	0.89	3.70	0.91	3.61	0.88	0.10
Surgency-extraversion	4.45	0.92	4.38	0.92	4.51	0.91	0.14
Effortful control	5.27	0.72	5.41	0.66	5.13	0.76	0.39

Note. *|d|*: Cohen's *d* coefficient (effect size).

Table 5.

Association between CBQ-dimensions and SDQ and BRIEF (Pearson correlations), and internal consistency of measures.

	Reliability		(n = 612)				
	external criteria		CBQ-SF		CBQ-VSF		
	Cronbach's	Negative	Surgency/	Effortful	Negative	Surgency/	Effortful
	alpha	affect	extraversion	control	affect	extraversion	control
SDQ-parents: Emotional symptoms	.52	.37*	-.11	-.07	.34*	-.16	-.01
SDQ-parents: Behavioral problems	.54	.17	.25	-.21	.16	.17	-.11
SDQ-parents: ADHD	.54	-.07	.51*	-.41*	-.07	.40*	-.27
SDQ-parents: Peers	.73	.18	-.05	-.12	.20	-.10	-.01
SDQ-parents: Pro-social	.55	.19	-.01	-.28	.18	-.03	-.21
SDQ-parents: Internalizing problems	.62	.33*	-.10	-.12	.33*	-.16	-.01
SDQ-parents: Externalizing problems	.72	.04	.48*	-.39*	.03	.37*	-.25
SDQ-parents: Total	.73	.20	.31*	-.35*	.19	.19	-.19
SDQ-teachers: Emotional symptoms	.69	.12	-.15	.01	.12	-.17	-.02
SDQ-teachers: Behavioral problems	.62	.03	.18	-.15	.00	.17	-.12
SDQ-teachers: ADHD	.64	-.10	.25	-.30*	-.09	.19	-.26
SDQ-teachers: Peers	.76	.09	-.09	-.12	.09	-.15	-.07
SDQ-teachers: Pro-social	.78	.02	.03	-.24	.03	.02	-.19
SDQ-teachers: Internalized problems	.72	.12	-.15	-.07	.13	-.19	-.06
SDQ-teachers: Externalized problems	.77	-.05	.26	-.28	-.06	.21	-.24
SDQ-teachers: Total	.80	.02	.12	-.24	.02	.06	-.20
BRIEF: Inhibitory Self-Control Index (ISCI)	.94	-.08	.20	-.23	-.08	.16	-.19
BRIEF: Flexibility Index (FI)	.91	.17	-.11	-.08	.16	-.14	-.07
BRIEF: Emergent Metacognition Index (EMI)	.96	-.01	.03	-.33*	.01	-.01	-.31*
BRIEF: Global Executive Composite (GEC)	.97	.00	.07	-.30*	.01	.03	-.26

Note. *Bold, correlations with good effect size ($|r| \geq .30$). Time stability valued with intraclass correlation coefficient (ICC). --- BRIEF questionnaire was not applied at age 4.

Table 6.

Association between CBQ-dimensions and DSM-IV disorders: OR coefficients and AUC in logistic regressions.

↓DSM-IV disorders	² CBQ and DSM-IV disorders at age 3 years (n = 612)								³ CBQ at age 3 and DSM-IV disorders at age 4 (n = 594)							
	CBQ-dimensions →			AUC	¹ Nega-			AUC	Nega-			AUC	¹ Nega-			AUC
tive	Sur-	Effort-	tive		Sur-	Effort-	tive		Sur-	Effort-	tive		Sur-	Effort-		
Any DSM-IV disorder	1.82*	1.54*	1.01	.653*	1.59*	1.31*	1.21	.634*	1.60*	1.28	0.75	.653*	1.57*	1.13	0.79	.647*
Any disruptive disorder	1.33	2.27*	0.27*	.782*	1.47*	1.99*	0.40*	.739*	1.96*	1.78*	0.57	.744*	1.85*	1.21	0.83	.693*
ADHD disorder	0.87	3.30*	0.13*	.876*	1.21	2.48*	0.24*	.814*	1.10	2.57*	0.48*	.808*	1.22	1.50	0.74	.708*
Oppositional defiant disorder	2.00*	1.61*	0.58*	.742*	1.74*	1.45*	0.72	.728*	2.45*	1.42	0.71	.724*	2.08*	1.11	1.07	.700*
Conduct disorder	0.60	1.79	0.38*	.778*	0.91	1.83	0.61*	.728*	---	---	---	---	---	---	---	---
Any anxiety disorder	3.78*	0.69	1.21	.755*	2.72*	0.73	1.42	.734*	1.86*	1.57	1.93*	.676*	1.91*	1.38	1.44	.690*
Separation anxiety	4.23*	0.95	1.04	.750*	2.98*	0.77	1.23	.737*	1.35	2.23*	1.63	.723*	1.76	1.94	1.19	.760*
Specific phobia	3.75*	1.17	1.47	.762*	2.83*	1.14	1.29	.749*	1.95*	2.08*	2.36*	.691*	1.77*	1.67	1.60	.683*
Social phobia	2.32*	0.29*	0.38	.849*	1.70	0.44*	0.68	.799*	1.96*	0.29*	0.77	.865*	2.26*	0.45*	0.57	.875*
Impairment: CGAS < 70	2.39*	1.28	0.42*	.728*	2.10*	1.15	0.67*	.685*	1.82*	1.42	0.75	.701*	1.91*	1.34	0.78	.690*

Note. ¹Scales for the VSF of the CBQ. ²Results adjusted for children's sex and other comorbidities. ³Results adjusted for children's sex, other comorbidities and the presence of the disorder at age 3.

AUC: area under the ROC curve. *Bold: significant result (.05 level).. --- Not estimable due to low prevalence of the disorder.